

# Artificial neural networks approach in evapotranspiration modeling: a review

M. Kumar · N. S. Raghuwanshi · R. Singh

Received: 6 January 2010 / Accepted: 22 July 2010 / Published online: 5 August 2010  
© Springer-Verlag 2010

**Abstract** The use of artificial neural networks (ANNs) in estimation of evapotranspiration has received enormous interest in the present decade. Several methodologies have been reported in the literature to realize the ANN modeling of evapotranspiration process. The present review discusses these methodologies including ANN architecture development, selection of training algorithm, and performance criteria. The paper also discusses the future research needs in ANN modeling of evapotranspiration to establish this methodology as an alternative to the existing methods of evapotranspiration estimation.

## Introduction

Reference evapotranspiration (ET<sub>o</sub>) is one of the major components of the hydrologic cycle, and its accurate estimation is of utmost importance for many studies such as hydrologic water balance, irrigation system design and management, crop yield simulation, and water resources planning and management. The ET<sub>o</sub> can be directly measured using either lysimeter or water balance approach, or estimated indirectly using the climatological data. However, it is not always possible to measure ET<sub>o</sub> using

lysimeter because it is a time-consuming method and needs precisely and carefully planned experiments. The indirect reference evapotranspiration (ET<sub>o</sub>) estimation methods based on climatological data vary from empirical relationships to combination methods based on physical processes such as Penman (1948) and Monteith (1965). The combination method links evaporation dynamics with fluxes of net radiation and aerodynamic transport characteristics of natural surface.

ET<sub>o</sub> is a complex process depending on several interacting climatological factors, such as temperature, humidity, wind speed, and radiation. The lack of physical understanding of ET<sub>o</sub> process and unavailability of all relevant data results in inaccurate estimation of ET<sub>o</sub>. Over the past two decades, artificial neural networks (ANNs) have been increasingly employed in modeling of hydrological processes because of their ability in mapping the input–output relationship without any understanding of physical process. An artificial neural network model is a mathematical model whose architecture is essentially analogous to the learning capability of the human brain in which highly interconnected processing elements arranged in layers. The basic strategy for developing an artificial neural network-based model of system behavior is to train the network on examples of the system. If the examples contain the relevant information about the system behavior, then the trained neural network would contain sufficient information about the system behavior to qualify as a system model. Such a trained neural network not only able to reproduce the results of examples it was trained on, but through its generalization capability, is able to approximate the results of other examples.

ANN application and its functional evaluation in the area of water resources can be found in the literature, those includes Dawson and Wilby (1998), Maier and Dandy

---

Communicated by S. Raine.

---

M. Kumar (✉)  
Division of Resource Management, Central Research Institute  
for Dryland Agriculture, Hyderabad, Andhra Pradesh 500 059,  
India  
e-mail: manovpkas@rediffmail.com

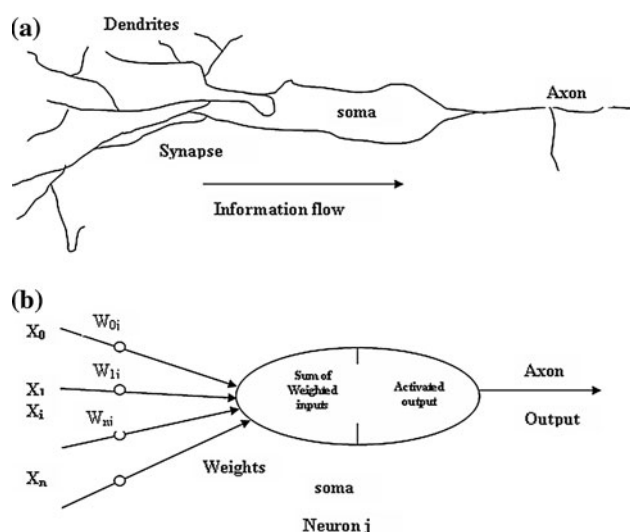
N. S. Raghuwanshi · R. Singh  
Agricultural and Food Engineering Department, Indian Institute  
of Technology, Kharagpur, West Bengal 721 302, India

(2000), and ASCE (2000a, b). The feed-forward multilayer perceptron (MLP) is widely adopted ANN in most of the studies on hydrological modeling. French et al. (1992) reported application of ANN in rainfall modeling whereas the application of ANN in evapotranspiration modeling was started after 2000 (Odhiambo et al. 2001; Kumar et al. 2002). Earlier, the ANN models for evapotranspiration used the standard back-propagation learning algorithm. Later, these models were further refined by using different learning algorithms such as radial basis function, RBF (Sudheer et al. 2003), generalized regression neural network model, GRNNM (Kisi 2006; Kim and Kim 2008), Levenberg–Marquardt, LM (Kisi 2007, 2008), quick propagation (Landeras et al. 2008), and conjugate gradient descent (Landeras et al. 2008). In the last decade, several studies on application of ANN techniques in evapotranspiration modeling have been reported considering different learning algorithms, input variables, training data sets, etc., and thus, this paper is an attempt to review these studies.

## Development of ANN as a computational technique

### Artificial neural networks: theory

An ANN consists of input, hidden, and output layers, and each layer includes an array of artificial neurons. A fully connected neural network, in which there is a connection between each of the neurons in any given layer with each of the neurons in the next layer or previous layer. An artificial neuron is a mathematical model whose components are analogous to the components of actual neuron. Figure 1 presents the various components of biological



**Fig. 1** Schematic representation of **a** A biological neuron, **b** an artificial neuron

neurons and its mathematical equivalence. The array of input parameters is stored in the input layer, and each input variable is represented by a neuron. Each of these inputs is modified by a weight (sometimes called synaptic weight) whose function is analogous to that of the synaptic junction in a biological neuron. The neuron (processing element) consists of two parts. The first part simply aggregates the weighted inputs resulting in a quantity  $I$  (weighted input); the second part is essentially a nonlinear filter, usually called the transfer function or activation function. The logistic and hyperbolic functions are the most common and have been used in majority of ANN application (Dawson and Wilby 1998 and Zanetti et al. 2007) in hydrological modeling. The activation function limits the values of the output of an artificial neuron to values between two asymptotes. These functions are a continuous function that varies gradually between two asymptotic values, typically 0 and 1 or  $-1$  and  $+1$ .

Considering a neuron  $j$  in the  $i$ th layer, sum of weighted inputs ( $I_j$ ) for a neuron  $j$  in the  $i$ th layer can be written as the dot product of the weight and input vectors as follows:

$$I_j = \sum_{i=1}^n w_{ij}x_i \quad (1)$$

where  $I_j$  = sum of weighted inputs to neuron  $j$  in the  $i$ th layer,  $w_{ij}$  = weight of link between the  $i$ th input to the  $j$ th neuron of the hidden layer,  $x_i$  = input value stored in  $i$ th neuron, and  $n$  = number of inputs.

The output of  $j$ th neuron can be written as a function of sum of weighted inputs:

$$y_i = \varphi(I_j) \quad (2)$$

where  $y_j$  = output of  $j$ th neuron and  $\varphi(I_j)$  is the activation function.

The activation function (logistic function) is given as

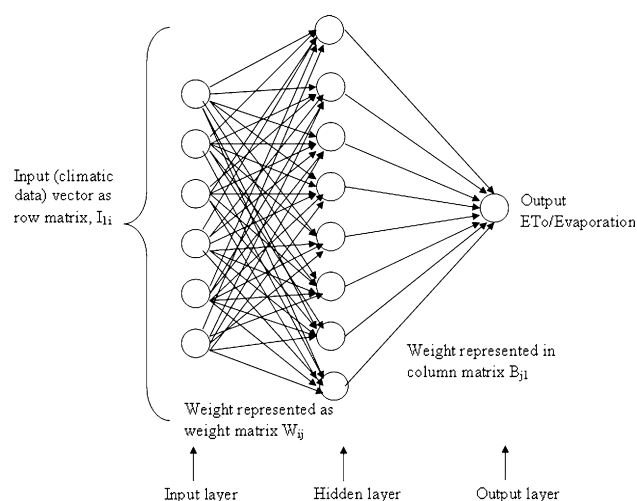
$$\varphi(I_j) = \frac{1}{1 + e^{-\alpha I_j}} \quad (3)$$

where  $\alpha$  = coefficient that adjusts the abruptness of the function and  $I_j$  = sum of the weighted input.

The hyperbolic tangent function is also utilized as an alternate to the logistic function. In fact, both the function forms sigmoid in nature. Mathematically, the hyperbolic tangent function is given as

$$\varphi(I_j) = \tanh(I) \quad (4)$$

The implementation of feed-forward ANN involves six different components that includes definition of input layers and number of nodes, selection of activation function, optimization of hidden layers and number of nodes, definition of output layer and number of nodes, selection of training and validation algorithm and performance evaluation. A typical ANN architecture consisting of an input



**Fig. 2** ANN frame work for ETo modeling

layer, a hidden layer and an output layer for modeling evapotranspiration is given in Fig. 2. The input layer consists of nodes representing meteorological inputs (namely minimum and maximum temperature, minimum and maximum RH, solar radiation and wind speed), and the output layer consists a node representing ETo.

#### Evapotranspiration modeling using ANN: past studies

The ANN applications in hydrological studies were started from rainfall-runoff modeling during the early 1990s. Later, the application further proliferated into streamflow prediction, ground water modeling, water quality, precipitation forecasting, reservoir operation and time series analysis of hydrological processes. The evapotranspiration modeling using ANN initiated after year 2000 when Bruton et al. (2000) developed artificial neural network models to estimate daily pan evaporation using weather data from Rome, Plains, and Watkinsville, Georgia. The ANN models estimated pan evaporation slightly better than the pan evaporation estimated using multiple linear regression models and Priestley–Taylor equation.

Kumar et al. (2002) developed ANN models using six basic parameters, which are also required for FAO-56 PM (Penman–Monteith) method. They trained ANNs against FAO-56 estimated ETo and lysimeter measured daily ETo for Davis. The ET was better estimated using the developed ANN models than the FAO-56 PM method despite ANN models are essentially empirical in nature. The WSEE (weighted standard error of estimate) was 0.56 mm/day for ANN model when compared to 0.97 mm/day in case of FAO-56 PM. Kisi (2006) developed ANN models respective to Penman (Snyder and Pruitt 1985), Hargreaves (Hargreaves and Samani 1985) and Ritchie (Jones and Ritchie 1990) for two weather stations namely Pomona and

Santa Monica, both operated by CIMIS (California Irrigation Management Information System). The ANN models were trained against the computed ETo estimates using FAO-56 PM. Different combination of weather parameters namely solar radiation, air temperature, relative humidity, and wind speed were used in developing the ANN models. The ANN model with all parameters as input was found to be the better estimator of daily FAO-56 PM than respective counterparts.

ANN model development in respect to temperature-based Hargreaves method has been given major attention by several researchers. This method is one of the simplest ETo estimation methods which requires maximum and minimum temperature as input parameters, and can be utilized over various regions. Conventional Hargreaves method over and under estimate ETo by 25 and 9%, respectively, in arid and humid region (Jensen et al. 1990). Khoob (2008a) developed ANN model corresponding to the Hargreaves method using monthly data of 12 locations situated in Khuzestan plain, Iran. The analysis of results suggested that the conventional Hargreaves method underestimates and overestimates the FAO-56 PM monthly ETo values by maximum of 20 and 37%, respectively. Khoob (2008b) concluded that the Hargreaves method is poor for regional estimation of ETo and calibrated Hargreaves method (using FAO-56PM) also does not improve ETo estimates. Kisi (2006) reported though Hargreaves method can overestimate FAO-56PM ETo by 118–167% but calibration significantly improves the performance of the temperature-based method. Zanetti et al. (2007) and Landaras et al. (2008) also found that for same inputs ANN models are better than the Hargreaves method. However, in contrast to these findings, Kisi (2006, 2007 and 2008) found the temperature-based models including Hargreaves when calibrated using FAO-56PM performed slightly better than ANN.

Kisi (2006) developed generalized regression neural networks (GRNN) model that uses radial basis function in ANN learning and tested at two locations, namely Pomona and Santa Monica. Two models were developed based on different data sets. The first model GRNN1 used data on solar radiation, temperature, relative humidity and wind speed whereas second model GRNN2 used solar radiation and temperature as model input. Inter-comparison of GRNN1 and GRNN2 suggested that the model performance did not improve substantially after adding input parameters of relative humidity and wind speed. Basic data of solar radiation and temperature was proved to be sufficient to estimate ETo. Landaras et al. (2008) developed series of ANN models to estimate ETo depending upon the availability of data. Kumar et al. (2008) developed ANN models based on different categories of conventional ETo estimation methods namely temperature based (FAO-24

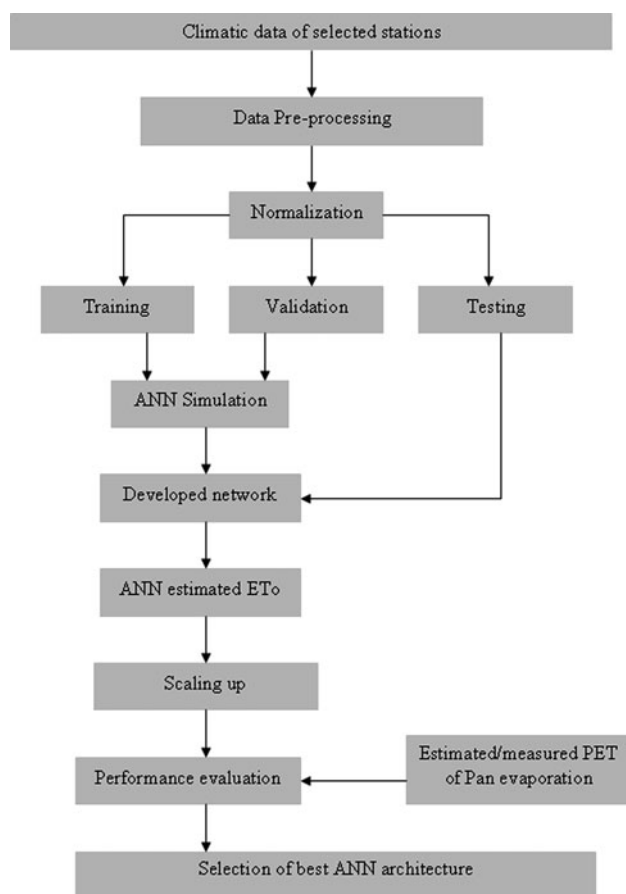
Blaney–Criddle), solar radiation based (FAO-24 Radiation and Turc) and combinations of these (FAO-56 Penman–Monteith). All ANN models performed better than the respective conventional method in estimating FAO-56PM

ETo. In general, all of the above studies revealed that the ANN models are superior in estimating ETo than the respective conventional method. Table 1 presents the summary of ANN-based ET modeling studies.

**Table 1** Summary of various criteria adopted by different authors in ANN modeling for ETo estimation

S No	Authors	Input data <sup>a</sup>	Output/target data	Length of record	Origin	Time step
1	Bruton et al. (2000)	$T_{\text{mean}}, R, T_{\text{soil}}, u, S_{\text{ra}},$ and $E$	Pan evaporation	7 years (1992–1998)	Georgia (USA)	Daily
2	Trajkovic et al. (2000)	$RH_{\text{min}}, n/N,$ and $u$	Blaney–Criddle ‘b’ factor	Tabular values of ‘b’ factor	–	–
3	Sudheer et al. (2002)	$T_{\text{min}}, T_{\text{max}}, RH_{\text{min}}, RH_{\text{max}}, n,$ and $u$	Pan evaporation	4 years (duration not mentioned)	India	Daily
4	Kumar et al. (2002)	$S_{\text{ra}}, T_{\text{min}}, T_{\text{max}}, RH_{\text{min}}, RH_{\text{max}},$ and $u$	ETo(FAO-56 PM) ETo(Lysimeter)	10 years (1991–2000) 4 years (1960–1963)	CIMIS (USA)	Daily
5	Sudheer et al. (2003)	$T_{\text{mean}}, RH_{\text{mean}}, u,$ and $n$	ETo(Lysimeter)	4 months (Oct 1989–Jan 1990)	India	Daily
6	Trajkovic et al. (2003)	$T_{\text{mean}}, RH_{\text{mean}}, u,$ and $n$	ETo(FAO-56 PM)	20 years (1977–1996)	Serbia	Monthly
7	Trajkovic (2005)	$T_{\text{mean}}, RH_{\text{mean}}, u, n,$ and $e_a$	ETo(FAO-56 PM)	4–8 years (different duration)	Serbia	Monthly
8	Keskin and Terzi (2006)	$T_{\text{mean}}, T_{\text{water}}, RH_{\text{mean}}, S_{\text{ra}}, u,$ and $n$	Pan evaporation	2 years (2001–2002)	Turkey	Daily
9	Kisi (2006)	$S_{\text{ra}}, T_{\text{mean}}, RH_{\text{mean}}$ and $u$	ETo(modified penman)	4 years (2001–2004)	CIMIS, USA	Daily
10	Kisi (2007)	$c, S_{\text{ra}}, T_{\text{mean}}, RH_{\text{mean}},$ and $u$	ETo(modified penman)	4 years (2001–2004)	CIMIS, USA	Daily
11	Kisi and Ozturk (2007)	$S_{\text{ra}}, T_{\text{mean}}, RH_{\text{mean}}$ and $u$	ETo(FAO-56 PM)	4 years (2001–2004)	CIMIS, USA	Daily
12	Kisi (2008)	$S_{\text{ra}}, T_{\text{mean}}, RH_{\text{mean}},$ and $u$	ETo(FAO-56 PM)	4 years (2001–2004)	CIMIS, USA	Daily
13	Jain et al. (2008)	$T_{\text{mean}}, T_{\text{dew}}, S_{\text{ra}}, u,$ and $RH_{\text{mean}}$	ETo(FAO-56 PM)	2 years (1990–1991)	USDA	Hourly
14	Landeras et al. (2008)	$T_{\text{mean}}, RH_{\text{mean}}, u,$ and $n$	ETo(FAO-56 PM)	5 years (1999–2003)	Spain	Daily
15	Zanetti et al. (2007)	$T_{\text{mean}}, S_{\text{ra}},$ and $n$	ETo(FAO-56 PM)	8 years (1996–2003)	Brazil	Daily
16	Khoob (2008a)	$T_{\text{mean}}, RH_{\text{mean}}, u$ and $n$	ETo(FAO-56 PM)	None revealed	Iran	Daily
17	Khoob (2008b)	$T_{\text{mean}}, RH_{\text{mean}}, u, n$ and $E$	ETo(FAO-56 PM)	8 years (1996–2003)	Iran	Daily
18	Kumar et al. (2008)	$S_{\text{ra}}, T_{\text{min}}, T_{\text{max}}, RH_{\text{min}}, RH_{\text{max}},$ and $u$	ETo(FAO-56 PM)	10 years (1991–2000)	CIMIS (USA)	Daily
19	Kumar et al. (2009)	$S_{\text{ra}}, T_{\text{min}}, T_{\text{max}}, RH_{\text{min}}, RH_{\text{max}},$ and $u$	ETo(FAO-56 PM)	10 years (1991–2000)	CIMIS (USA)	Daily
20	Kisi and Cimen (2009)	$S_{\text{ra}}, T_{\text{mean}}, RH_{\text{mean}},$ and $u$	ETo(FAO-56 PM)	10 years (1998–2007)	CIMIS, USA	Daily
21	Kisi (2009)	$T_{\text{mean}}, S_{\text{ra}}, RH,$ and $e_{\text{air}}$	Pan evaporation	4 years (1987–1990)	Fresno, Los Angeles and San Diego, USA	Daily
22	Dai et al. (2009)	$T_{\text{mean}}, RH, u,$ and $n$	ETo(FAO-56 PM)	30 years (1970–2000)	China	Monthly
23	Moghaddamnia et al. (2009)	$T_{\text{mean}}, u, e_d,$ and $RH$	Pan evaporation	23 years (1983–2005)	Iran	Daily
24	Traore et al. (2010)	$S_{\text{ra}}, T_{\text{min}}, T_{\text{max}}, RH,$ and $u$	ETo(FAO-56 PM)	11 years (1996–2003)	Sudano-Sahelian Zone	Daily
25	Marti et al. (2010)	$T_{\text{mean}}, T_{\text{min}}, T_{\text{max}}, RH_{\text{min}}, RH_{\text{max}}, u, S_{\text{ra}},$ and $n$	ETo(FAO-56 PM)	9 years (2000–2008)	Spain	Daily

<sup>a</sup> The symbols are  $T_{\text{mean}}$  air temperature,  $T_{\text{min}}$  minimum temperature,  $T_{\text{max}}$  maximum temperature,  $T_{\text{dew}}$  dew point temperature,  $R$  rainfall,  $T_{\text{soil}}$  soil temperature,  $u$  wind speed,  $S_{\text{ra}}$  solar radiation,  $E$  pan evaporation,  $RH_{\text{min}}$  minimum  $RH$ ,  $RH_{\text{max}}$  maximum  $RH$ ,  $n$  actual sunshine hours,  $N$  maximum possible sunshine hours,  $e_a$  actual vapor pressure, and  $e_{\text{air}}$  air pressure,  $e_d$  saturation vapor pressure deficit



**Fig. 3** Methodology for developing ANN architecture

Kumar et al. (2009) developed generalized artificial neural network (GANN) models, namely GANN(6-7-1), GANN(3-4-1), GANN(5-6-1) and GANN(4-5-1) corresponding to FAO-56PM, FAO-24 Radiation, Turc and FAO-24 Blaney–Criddle method. These models were trained using the pooled data from 4 CIMIS stations namely Castroville, Davis, Mulberry, and WSFS with FAO-56 PM computed values as targets. The developed GANN models were then directly implemented at two different CIMIS locations in California, USA (namely Lodi and Fresno) and four IMD (India Meteorological Department) locations in Madhya Pradesh and Chattisgarh state of India (namely Hoshangabad, Gwalior, Jabalpur and Pendra) without any local training. Based on the results they concluded that the developed GANN models have the global validity and can be used directly to estimate ETo under arid conditions.

The general methodology adopted in implementing the ANN techniques for evapotranspiration modeling is comparison of ANN model performance with the respective conventional method. The framework of these studies can be summarized as Fig. 3. The methodology essentially looks into the possibility of use of ANN model as an

alternative to the respective conventional method. Several studies have focused on developing ANN models corresponding to various conventional ETo estimation methods.

#### Development of ANN architecture: some issues

ANN architecture development involves the creation of network topology and network training under various combinations of nodes in hidden layer(s), number of hidden layers, training cycles, and selection of training function and parameters of training function. The performance of each combination is evaluated on the basis of predefined performance criteria. Typical ANN model has three layers namely input layer, hidden layer and output layer having certain number of nodes in each layer (Fig. 2). Nodes in input layer and output layer are known, and they typically represent input variables and expected output, respectively.

#### Nodes in different layers of ANN models

The number of nodes in the input layer depends on the number of climatic variables used in estimating ETo. The individual node in the input layer corresponds to respective variables. Thus, the number of nodes in the input layer varies according to the climatic data requirement of the model. The number of hidden nodes depends on several factors such as the number of input and output nodes; the number of training cases; the amount of noise in the targets; the complexity of the function or classification to be learned; the architecture; the type of hidden unit activation function; and the training algorithm. In most situations, there is no way to determine the optimal number of hidden nodes without training several networks and estimating the error of each. A few hidden nodes result in a high training and generalization errors due to under-fitting, whereas too many hidden nodes result in low training error but give higher estimation error due to over-fitting. However, Swingler (1996) and Berry and Linoff (1997) pointed out that the networks rarely require hidden nodes more than twice the number of input nodes for the MLP.

Kumar et al. (2002) studied the effect of number of nodes in the hidden layer and number of hidden layers on the evapotranspiration estimating performance. In their study, the number of nodes was varied from  $i + 1$  to  $2i$  (where  $i$  is the number of nodes in the input layer) in two sets of ANN models, i.e., (a) ANN with single hidden layer and (b) ANN with two hidden layers. The study revealed that a single hidden layer with  $i + 1$  nodes was sufficient to model evapotranspiration for almost all conditions of data availability. Jain et al. (2008) also adopted the trial and error procedure to optimize the number of nodes in the hidden layer as well as number of hidden layer(s). The procedure started with 2 nodes in the hidden layer initially



and number of nodes in the hidden layer increased to 12 with step size of 1 at each trial. Several researchers namely (Sudheer et al. 2003; Kisi 2007, 2008; Landeras et al. 2008) have used the in-built utilities of standard ANN simulation software to achieve the optimal ANN architecture. These in-built utilities of simulation software usually adopt network pruning algorithm. Initially bigger ANN architecture is selected and simulation implemented with this ANN. The pruning algorithm reduces the network size by removing the links with the smallest weight after each training cycle. The pruning of network, however, has not been reported explicitly in the literature though. Trajkovic et al. (2003) used optimal brain surgeon (OBS) and optimal brain damage (OBD) algorithms for pruning to derive an optimal network architecture based on the statistical parameter significant test. The method of finding an optimal network architecture using a pruning algorithm has certain drawbacks. This type of ANN usually function for certain sets of conditions and parameters specific to the location. However, the network parameters may vary with respect to location and because of variability in input data. Thus, the optimal network architecture obtained using pruning algorithm will have limited scope to generalize over various locations.

#### Selection of learning function

The goal of neural network learning is to find the network having the best performance on unseen data. Learning algorithm is the mathematical formulation that optimizes the error function in order to modify the link weight. Most of the studies on ANN modeling of hydrological events used the back-propagation. In evapotranspiration modeling, varieties of learning algorithms are employed in ANN model development. These algorithms are back-propagation (Kumar et al. 2002; Sudheer et al. 2002; Kisi 2006; Keskin and Terzi 2006; Parsuraman et al. 2007; Khoob 2008a; Landeras et al. 2008; Kumar et al. 2008), generalized regression (Kisi 2006, 2008; Kim and Kim 2008), radial basis function (Sudheer et al. 2003; Trajkovic et al. 2003; Trajkovic 2005; Kisi 2008), Levenberg-Maquardt (Kisi 2007; Landeras et al. 2008; Zanetti et al. 2007; Khoob 2008b), conjugate gradient decent (Landaras et al. 2008), and quick propagation (Landaras et al. 2008).

#### Back-propagation algorithm

More than 70% of the studies on ANN implementation of hydrological processes employed the back-propagation learning algorithm because of its simplicity and robustness. In the back-propagation learning algorithm, learning of an artificial neural network involves two phases. In the forward pass the input signals propagate from the network

input to the output. In the reverse pass, the calculated error signals propagate backward through the network, where these are used to adjust the weights (Rumelhart et al. 1986). The calculation of the output is carried out, layer by layer, in the forward direction. The output of one layer is the input to the next layer. In the reverse pass, the weights of the output neuron layer are adjusted first since the target value of each output neuron is available to guide the adjustment of the associated weights, using the *delta rule*. Next, the weights of middle layer (hidden layer) are adjusted. Lower learning rate results in slow learning of the networks, whereas higher training rate makes the weights and objective (error) function diverge, so there is no training at all (LeCun et al. 1989, 1993; Orr and Leen 1997). Therefore, network training using a constant training rate is usually finalized after trial and error. Use of momentum term reduces the training time by enhancing the stability of the training process and keeps the training process going in the same general direction. This involves adding a term to the weight adjustment function, which is proportional to the previous weight change to modify the next weight change to overcome the ill effects of local minima. The momentum term carries a weight change process through one or more local minima and gets it into the global minima. ASCE Task Committee (2000a) described the algorithm in detail.

Initial weights are assigned to each and every link between the processing elements of different layers to start the learning process of the network. Generally, initial weights are assigned randomly, hence there is a little chance that the two same networks with same input data will converge to some final state. However, this aspect of ANN learning is automatically taken care when ANN is learned for predetermined error values. Kumar et al. (2002, 2009) reported that ANN model of same architecture with different learning parameters did not produce significantly different results. This suggested that the variation to the learning rate, momentum, and initial weight range had very little role in model accuracy when sufficient data are available and predefined ANN learning criteria are used for training.

#### Radial basis function

The radial basis function networks (RBF networks) have been extensively employed in modeling of evapotranspiration process. Radial basis functions (RBF) are built onto distance criteria with respect to center. The interpretation of its output layer value is the same as a regression model in statistics. The RBF network consists of hidden layer of basis function (most commonly used is Gaussian Bell function) represented as hidden layer. At the input of each neuron, the distance between the neuron center in hidden

layer and input vector is calculated and the output is determined by summation of dot products of basis function and distance with respect to each input neuron. Thus, the RBF networks have an advantage of not suffering from the problem of local minima as in the case of the back-propagation MLP. The linearity ensures that the error surface is quadratic and therefore a global minimum is found easily. The principle of radial basis function is derived from the theory of functional approximation. The RBF function is presented in details below (SNNS 1995). Given  $N$  pairs  $(\vec{x}_i, y_i)$  and  $(\vec{x} \in \mathbb{R}^n, y \in \mathbb{R})$ , the output function of the network is

$$\int(\vec{x}) = \sum_{i=1}^K c_i h(|\vec{x} - \vec{t}_i|) \quad (5)$$

where  $K$  is the number of neurons in the hidden layer,  $\vec{t}_i$  is the center vector for neuron  $i$ , and  $c_i$  are the weights of the linear output neurons. In the basic form all inputs are connected to each hidden neuron.  $h$  is the radial basis function applied to the Euclidian distance between each center  $\vec{t}_i$  and the given argument  $\vec{x}$ . The function  $h$  is basically a Gaussian function which has its maximum value at a distance of zero. When values of  $\vec{x}$  equals to center,  $\vec{t}_i$ , the function  $h$  yields output value of 1 and become zero for larger distances. Hence the function  $\int$  is an approximation of the  $N$  given pairs  $(\vec{x}_i, y_i)$  and therefore minimizes the following error function,  $H$ :

$$H[f] = \sum (y_i - f(\vec{x}_i))^2 + \lambda \|Pf\|^2 \quad (6)$$

The first part of the function,  $H$ , is the condition which minimizes the local error of the approximation and the second part of  $H$ , i.e.,  $\|Pf\|^2$  is a stabilizer which forces  $f$  to become continuous. The factor  $\lambda$  determines the influence of the stabilizer.

Sudheer et al. (2003) used RBF function to train the networks for actual ET estimation using three different input datasets corresponding to FAO-56, Hargreaves, and Thornthwaite methods. These developed networks could estimate actual ET with modeling efficiency of 98.2–99.0%. Generalized regression neural networks (GRNN) models are extension to the RBF networks that use one additional layer known as summation layer. The GRNN models are based on nonlinear regression theory. The GRNN thus, serves as the universal approximator in solving any smooth function approximation problem. The method does not require iterative learning process. Hence, local minima problem is not faced in GRNN modeling. The detailed of the learning procedure of GRNN models is presented by Kisi (2006) and Kim and Kim (2008). Trajkovic (2005) demonstrated the utility of a temperature-based radial basis function neural networks in modeling ETo when relative humidity, solar radiation and

wind speed data are not available. Kisi (2006) applied GRNN technique in modeling FAO-56 PM ETo and also compared its performance against Penman, Hargreaves, and Ritchie methods. Trajkovic (2009) further used RBF networks to convert pan evaporation data into evapotranspiration.

#### Levenberg–Marquardt (LM)

During past couple of years, the Levenberg–Marquardt (LM), a second order optimization technique is extensively employed in evapotranspiration modeling using artificial neural networks. LM algorithm is modification to the Gauss–Newton method of approximation and presented in detail by Kisi (2007). The LM algorithm is claimed to be superior to the standard back-propagation in terms of fast convergence and thus needs lesser learning cycles (Zanetti et al. 2007).

#### Data type, quality, and sample size

In general, all the studies of evapotranspiration modeling using ANN involved comparison of the traditional evapotranspiration estimation methods and counterpart ANN models. Data such as temperature, humidity, wind speed, and solar radiation are used in ANN modeling of evapotranspiration. In most of the studies, the target values of evapotranspiration derived from the traditional methods, based on temperature or radiation or combination methods, were used in respective ANN model development and testing. However, Kumar et al. (2002) used lysimeter measured evapotranspiration for ANN model development. Bruton et al. (2000), developed ANN models to estimate missing pan evaporation data using weather variables either measured or calculated. The data used by different authors in the literature for ETo estimation using ANN modeling is described in Table 1. The basic data include solar radiation, humidity, temperature, and wind speed. ANN models were developed using different combination of these basic data.

The sample size requirement of input data pairs for optimal ANN learning has not been studied extensively in ETo modeling. However, Markham and Rakes (1998) carried out an exclusive study on the effect of sample size and data variability on performance of ANN (three layer back-propagation network with sigmoidal transfer function). The study involved 5 level of sample size viz. 20, 50, 100, 200 and 500 and 4 level of variance viz. 25, 100, 225 and 400. The study suggested that the sample size and variance have an effect on relative performance of ANN and ANN performs better with large sample size.

ANN models are data intensive and learn through data variability. A high variability in learning data set results in

a more robust ANN model. Input data should contain the values in wider range because the ANN has strong interpolation capabilities than the extrapolation capabilities (Kumar et al. 2009). Thus, ANNs performance deteriorates substantially in respect to the values which are beyond the range of input data on which network has been trained. Several studies suggested that the ANN modeling techniques has a close resemblance with multiple regression data modeling. Thus the sample size determination using the standard statistical methods in regression modeling could be employed in ANN modeling as well. These statistical methods consider three factors namely degree of confidence, degree of precision, and variability in data set in determining the optimal sample size (Cochran and Cox 1957; Gomez and Gomez 1984).

#### Data preprocessing (normalization)

Neural networks are very sensitive to absolute magnitude. To minimize the influence of absolute scale, all inputs to the neural network are scaled and normalized so that they correspond to similar range of values. The most commonly chosen ranges are 0 to 1 or  $-1$  to  $+1$ . This is described in the following section.

The nonlinear activation function used in network learning is typical sigmoidal function, which is given as

$$\Phi(I) = (1 + e^{-xI})^{-1} \quad (7)$$

where  $I$  is the sum of weighted inputs and given as

$$I = x_1w_1 + x_2w_2 + \cdots + x_nw_n = \sum_{i=1}^n x_iw_i \quad (8)$$

where  $x_i$  and  $w_i$  represent the element of input vector and associated weight, respectively.

This logistic function  $\Phi(I)$  is one of the several sigmoidal functions, in which values change monotonically from a lower limit (0 or  $-1$ ) to an upper limit ( $+1$ ) as  $I$  changes.

Prior to exporting the data to the ANN for training, normalization of the data is essential to restrict the data range within an interval of 0–1, because the sigmoidal activation function is assigned to the neurons of the middle layer. The weight change corresponding to the activated value near 0 or 1 is minimal since neuron response is “dull”, whereas closer to 0.5, the neurons respond more (Rao and Rao 1996).

Most of the researchers (Hegazy and Ayed 1998; Campolo et al. 1999; Yeh 1998) followed the normalization schemes such that the maximum of the data series fixed at unity and thus used maximum value to divide the data series.

$$x_{norm} = \frac{x_o}{x_{max}} \quad (9)$$

where  $x_{norm}$ ,  $x_o$ , and  $x_{max}$  are normalized value, real value and maximum value, respectively.

Statistical characteristics of the original data were used by some researchers to normalize the values. Khoob (2008a) utilized mean and standard deviation of the original data in normalization process. The normalization scheme given was

$$x_{norm} = \frac{x_o - \bar{x}}{SD} \quad (10)$$

where  $x_{norm}$ ,  $x_o$ ,  $\bar{x}$ , and  $SD$  are normalized value, real value, mean, and standard deviation, respectively.

Several researchers (Jain et al. 2008; Khoob 2008b) normalized the data between 0 and 1 using maximum and minimum of the observed data set as follows:

$$x_{norm} = \frac{x_o - x_{min}}{x_{max} - x_{min}} \quad (11)$$

where  $x_{norm}$ ,  $x_o$ ,  $x_{min}$ , and  $x_{max}$  are normalized value, real value, minimum value, and maximum value, respectively.

The shape of the transfer function plays an important role in ANN learning. The majority of neural network models use a sigmoid (S-shaped) transfer function. The sigmoid functions are defined as a nonlinear function whose first derivative is always positive and range is bounded. However, the sigmoid function never reaches to its theoretical maximum and minimum limits and hence considered active in the range of 0.1–0.9 (Rao and Rao 1996). Cigizoglu (2003) suggested that the learning data that are scaled between 0.2 and 0.8 gives the ANN, the flexibility to predict beyond the training range. Kisi (2009) suggested following equation to normalize data between the desired range  $a$  and  $b$

$$x_{norm} = a \frac{x_i - x_{min}}{x_{max} - x_{min}} + b \quad (12)$$

where  $x_{norm}$ ,  $x_o$ ,  $x_{min}$ , and  $x_{max}$  are normalized value, real value, minimum value and maximum value respectively and  $a$  and  $b$  are the scaling factor.

Kumar et al. (2002, 2008 and 2009) used the normalized scheme such that the difference between the original data and mean was divided by the range of the data series. This resulted in the normalized values ranging between  $-1$  to  $+1$  with zero mean. To obtain the mean of 0.5 of the normalized series, the axis was transformed to  $-1$  and the resultant was divided by two. Thus, the normalization was made ensuring that the mean of the data series should have the normalized value closer to 0.5 and in the process most of data were fall in range of 0.3–0.7 ensuring maximum response from the neuron.



$$x_{norm} = 0.5 \times \left( \frac{x_o - \bar{x}}{x_{\max} - x_{\min}} \right) + 0.5 \quad (13)$$

where  $x_{norm}$ ,  $x_o$ ,  $x_{\min}$ , and  $x_{\max}$  are normalized value, real value, minimum value, and maximum value, respectively.

#### Network performance criteria

At the beginning of the ANN model development process, it is important to clearly define the criteria to assess the performance of the developed model. The performance criteria have the significant impact on the optimal network architecture selection. The performance criteria decided on the basis of either system error that occurred during learning or prediction accuracy. The prediction accuracy generally measures the ANN ability in estimation using data that have not been utilized in the training process. Several performance criteria have been used in ANN modeling of evapotranspiration. Various performance measures used in ANN modeling of ETo are given in Table 2. In most of the studies, the ANN models were developed as an alternative to the FAO-56PM. However, ANN models were also developed pertaining to various cases of data availability and compared with the respective conventional ETo estimation methods. All studies validated the application of ANN modeling as an alternative to the existing ETo estimation methods.

#### Discussion

ANNs have been successfully implemented to model evapotranspiration in several studies. These studies indicated that the ANN models can be used as an alternative method to estimate ETo. The performance of ANN models reported in these studies was superior to respective conventional methods of ETo estimation.

Most of the studies considered ANN modeling of FAO-56 PM ETo values. The FAO-56 PM ETo values are the calculated values which were essentially derived from the basic climatic data. Using these calculated values as target, the high model performance of the ANN models could be attained. The model efficiency deteriorates when the target value of pan evaporation and Lysimeter ETo are used. The interaction of different factors influencing the process of evapotranspiration and evaporation is nonlinear in nature. However, more nonlinearity exists in evapotranspiration process because of transpiration components. Humidity has the most effect on evaporation followed by solar radiation and wind speed (Kisi 2009). The evapotranspiration process, however, influenced by six basic factors namely solar radiation, maximum and minimum temperature, maximum and minimum RH, and wind speed. On the basis of different

results reported in various studies, it was observed that the ANN attained high model performance for calculated values of FAO-56 PM followed by pan evaporation data. The ANN performance substantially reduced in case of lysimeter data. Kumar et al. (2002) obtained high model efficiency (WSEE value equal to 0.27 mm day<sup>-1</sup>) for ANN models with FAO-56 PM ETo as target. The WSEE of the same ANN architecture was reduced to 0.56 mm day<sup>-1</sup> when lysimeter ETo was used as the target value. This study suggested that the ANN learning is dependent on the extent of nonlinearity exists in the process that is modeled. Hence, more complex ANN models are required to model either pan evaporation or lysimeter measured ET/actual ET.

In many respect, ANN models are black-box models that perform one to one modeling and do not describe the physical relationship among the different components of the process. Evapotranspiration models, such as Penman–Monteith, describe the evapotranspiration in the framework of mathematical model of energy balance representing the understanding of physical process. The model is supplemented with certain assumptions and coefficients in respect to the portion of the process that is not well understood. In contrast to this, the ANNs are trained with input–output data pairs and in the process of learning it mimics the underlying evapotranspiration process where knowledge stored in the link weights and thresholds. Unfortunately, these link weight and threshold has not been studied for physical interpretation explicitly. However, Jain et al. (2008) examined the link weights for physical interpretation of 6 ANN models and determined the relative importance of the input variables in evapotranspiration modeling. One of the major problems with the ANN modeling is that it does not reveal the set of optimal weights, where underlying process is defined, after training. Hence, an ANN model developed for one location can not be implemented to other location without local training. Thus, ANN modeling can not replace the traditional modeling techniques in its present form of application.

#### Development of generalized ANN models as a global predictor of evapotranspiration

Perhaps, this is the limitation of ANN modeling in the present context of research. The ANN models are location specific and influenced by the variance in the input data whereas Penman–Monteith method based on energy balance are adopted in all region. The ANNs are good to interpolate but its capabilities in extrapolation is always under question. The possible approach to develop generalized ANN models is to increase the range of variation in the input data set. This can be achieved by incorporating data from several distinct locations representing different regions and conditions in model development. In this

**Table 2** Summary of studies carried on ANN modeling for ETo estimation from leading contributions

Author(s)	Best network	Corresponding ETo estimation methods	Length of data		Performance criteria and its value	Applicability of models to other stations
			Training & validation	Testing		
Bruton et al. (2000)	14-na <sup>a</sup> -1	Pan evaporation	1533 patterns	511 patterns	RMSE = 1.11 mm day <sup>-1</sup>	The ANN model was developed using data from 40 automated weather stations
Trajkovic et al. (2000)	Not mentioned	FAO Blaney–Criddle 'F' factor	186 sample of tabular 'b' values	20 sample of tabular 'b' values	MARE = 0.34%; MXARE = 1.8%; $R^2 = 1.0$	Not studied
Sudheer et al. (2002)	6-12-1	Pan evaporation	1991 (daily)	1990, 1992–1993 (daily)	RMSE = 1.15 to 2.01 mm day <sup>-1</sup> ; % error = -6.9 to 4.89%; Efficiency = 22.94 to 66.85%; $r = 0.55$ to 0.84	Not studied
Kumar et al. (2002)	6-7-1	FAO-56PM	Jan 1990–Jun 1996 (daily)	Jul 1996–Jun 2000 (daily)	WSEE = 0.29 mm day <sup>-1</sup>	Not studied
	6-7-1	Lysimeter ETo	Jan 1960–Dec 1960 (daily)	Jan 1961–Dec 1961 (daily)	WSEE = 0.56 mm day <sup>-1</sup>	
Sudheer et al. (2003)	1-10-1 (average temp. as input data)	FAO-56PM	Oct–Nov 1989 (daily)	Dec 1989–Jan 1990 (daily)	SEE = 0.03 mm day <sup>-1</sup> RMSE = 0.03 mm day <sup>-1</sup> Efficiency = 99%	Not studied
Trajkovic et al. (2003)	Not mentioned	FAO-56 PM	20 years (monthly)	16 years (monthly)	MSE = 131.13 mm <sup>2</sup>  MAE = 8.90 mm $R = 0.97$	Not studied
Trajkovic (2005)	Not mentioned	FAO-56 PM Hargreaves Thornthwaite	276 records (monthly)	240 records (monthly)	MXE = 0.49 to 0.85 mm day <sup>-1</sup>	Network was developed using pooled data of three stations and was tested on 5 stations that were not included during network training
Keskin and Terzi (2006)	3-6-1	Penman	Mar 1–Oct 31, 01	Mar 1–Oct 31, 02	MAE = 0.13 to 0.19 mm day <sup>-1</sup> RMSE = 0.16 to 0.27 mm day <sup>-1</sup> MSE = 0.007 mm <sup>2</sup> day <sup>-2</sup>	
Kisi (2006)	Not mentioned	FAO-56 PM	2001–2003 (daily)	2004 (daily)	$R^2 = 0.77$ MSE = 0.058 mm <sup>2</sup> day <sup>-2</sup> MAE = 0.184 mm day <sup>-1</sup> $R^2 = 0.985$	Not studied
Kisi (2007)	4-6-1	Penman	2001–2003 (daily)	2004 (daily)	MSE = 0.03 to 0.32 mm <sup>2</sup> day <sup>-2</sup>	Developed networks were tested for three stations after local training
					MAE = 0.14 to 0.26 mm day <sup>-1</sup>	

Table 2 continued

Author(s)	Best network	Corresponding ETo estimation methods	Length of data		Performance criteria and its value	Applicability of models to other stations
			Training & validation	Testing		
Kisi and Ozturk (2007)	2-3-1	Hargreaves and Ture	2001–2003 (daily)	2004 (daily)	$R^2 = 0.866$ to $0.992$ $MSE = 0.07$ to $0.38 \text{ mm}^2 \text{ day}^{-2}$ $MAE = 0.20$ to $0.33 \text{ mm day}^{-1}$ $R^2 = 0.84$ to $0.98$ $MAE = 0.18 \text{ mm day}^{-1}$	Developed ANN models were applied to other station
	4-8-1	CIMIS Penman, Hargreaves and Ritchie	2001–2003 (daily)	2004 (daily)		
	4-na-1	Penman; Hargreaves and Ritchie	2001–2003 (daily)	2004 (daily)	$R^2 = 0.969$ $MSE = 0.036 \text{ mm}^2 \text{ day}^{-2}$	Not studied
	5-8-1	FAO-56 PM	Feb 1990–Dec 1990	Feb 1991–Dec 1991	$MAE = 0.144 \text{ mm day}^{-1}$ $R^2 = 0.991$ $SEE = 0.26 \text{ mm day}^{-1}$ $AARE = 0.39 \text{ mm day}^{-1}$ $MAE = 0.09 \text{ mm day}^{-1}$ $MBE = -0.047$ to $0.012 \text{ mm day}^{-1}$	Not studied
Landaras et al. (2008)	Several networks (4-4-1 to 6-14-1)	FAO-56PM Hargreaves Priestley–Taylor Makkink and Turc	1999 to 2000	2002–2003		Not studied
	4-10-1	FAO-56PM	Sep 1996–Aug 2002	Sep 2002–Aug 2003	$MAE = 0.174$ to $0.0385 \text{ mm day}^{-1}$ $RMSE = 0.238$ to $0.531 \text{ mm day}^{-1}$ $MSE = 0.36$ to $0.43 \text{ mm}^2 \text{ day}^{-2}$ $R^2 = 0.81$ to $0.84$ $RMSE = 0.59 \text{ mm day}^{-1}$	Not studied
Zanetti et al. (2007)	4-10-1	FAO-56PM	Sep 1996–Aug 2002	Sep 2002–Aug 2003		The developed ANN model was tested over 6 stations
Khoob (2008a)	5-9-1	FAO-56PM	70% of 576 monthly data pairs	30% of 576 monthly data pairs	$MBE = 0.06 \text{ mm day}^{-1}$ $R^2 = 0.97$ $RMSE = 0.8 \text{ mm day}^{-1}$ $R^2 = 0.91$ $WSEE = 0.66 \text{ mm day}^{-1}$ $R^2 = 0.79$ $WSEE = 0.64 \text{ mm day}^{-1}$ $R^2 = 0.88$	
Khoob (2008b)	Not mentioned	Pan evaporation	1996–2001	2002–2003		Not studied
Kumar et al. (2008)	4-5-1	FAO-24 BC	Jan 1990–Jun 1996	Jul 1996–Jun 2000		Not tested
	3-4-1	FAO-24RAD	Jan 1990–Jun 1996	Jul 1996–Jun 2000		

Table 2 continued

Author(s)	Best network	Corresponding ETo estimation methods	Length of data		Performance criteria and its value	Applicability of models to other stations
			Training & validation	Testing		
Kumar et al. (2009)	5-6-1	Turc	Jan 1990–Jun 1996	Jul 1996–Jun 2000	WSEE = 0.51 mm day <sup>-1</sup> $R^2 = 0.88$	ANN models were tested for stations that were not included in ANN training. The results suggested that ANN model can be directly used to estimate ETo (FAO-56PM)
	4-5-1	FAO-24 BC	Jan 1990–Jun 1996	Jul 1996–Jun 2000	WSEE = 0.68 mm day <sup>-1</sup>	
	3-4-1	FAO-24RAD	Jan 1990–Jun 1996	Jul 1996–Jun 2000	$R^2 = 0.88$ WSEE = 0.71 mm day <sup>-1</sup>	
	5-6-1	Turc	Jan 1990–Jun 1996	Jul 1996–Jun 2000	$R^2 = 0.88$ WSEE = 0.65	
	6-7-1	FAO-56PM	Jan 1990–Jun 1996	Jul 1996–Jun 2000	$R^2 = 0.90$ WSEE = 0.46 mm day <sup>-1</sup>	
Kisi and Cimen (2009)	4-8-1	CIMIS Penman, Hargreaves, Ritchie and Turc	1998–2004	2005–2007	$R^2 = 0.94$ MAE = 0.10 mm day <sup>-1</sup>	Not studied
	Several networks (1-1-1 to 5-7-1)	Pan Evaporation	3 years (1987–1989)	1 year (1990)	$R^2 = 0.996$ RMSE = 0.2 to 1.18 mm day <sup>-1</sup>	Not studied
Kisi (2009)	Several networks (1-1-1 to 5-7-1)	Pan Evaporation	3 years (1987–89)	1 year (1990)	MAE = 0.14 to 0.76 mm day <sup>-1</sup> $R^2 = 0.7$ to 0.998 RE = -6.8 to 0.8% RMSE = 0.2 to 1.65 mm day <sup>-1</sup>	Not studied
	4-8-1	FAO-56PM	30 years (1970–2000)	30 years (1970–2000)	MAE = 0.14 to 1.22% $R^2 = 0.45$ to 0.998 RE = -6.8 to 0.8% RMSE = 0.223 mm day <sup>-1</sup>	Out of 135 stations, the network was trained using data from two-third stations and was tested over remaining one-third stations
					RE = 4.2% $R^2 = 0.987$	

**Table 2** continued

Author(s)	Best network	Corresponding ETo estimation methods	Length of data		Performance criteria and its value	Applicability of models to other stations
			Training & validation	Testing		
Moghaddamnia et al. (2009)	3-9-1	FAO-56PM	30 years (1970–2000)	30 years (1970–2000)	RMSE = 0.438 mm day <sup>-1</sup> RE = 9.5% $R^2 = 0.987$ RMSE = 1.51%	
	Not mentioned	Pan Evaporation	Not mentioned	Not mentioned		Not studied
					$R^2 = 0.97$	
Traore et al. (2010)	Not mentioned	FAO-56PM	01/01/96–12/31/2003	01/01/2006–12/31/2006	RMSE = 0.048 mm day <sup>-1</sup> MAE = 0.033 mm day <sup>-1</sup> $R^2 = 0.998$	Not studied
					MSE = 0.21 to 0.71 mm <sup>2</sup> day <sup>-2</sup>	
Marti et al. (2010)	Not mentioned	FAO-56PM	6 years (2000–2006)	2 years (2007–2008)		A generalized model was developed using climatic data of 9 stations and tested over 43 stations
					AARE = 0.13 to 0.21	

<sup>a</sup> Nodes in hidden layer were not available

The symbol represents as *SSE* standard error of estimate mm day<sup>-1</sup>, *SSE* sum of square error, *CC* correlation coefficient, *RMSE* root mean square error, *RE* relative error, *MAE* mean absolute error, *AARE* average absolute relative error, *MSE* mean square error, *WSEE* weighted standard error of estimate, mm day<sup>-1</sup>, *MBE* mean bias error,  $R^2$  Nash–Sutcliffe coefficient, and *MARE* mean absolute relative error, *MXARE* maximum absolute relative error



situation, however, ANN probably will have lesser accuracy than ANN models with local training. Kumar et al. (2009) pooled the data of four distinct locations in ANN learning to develop generalized ANN models having basic variables identical to the Penman–Monteith. The ANN model performed satisfactory for arid location but performance deteriorated in humid location due to under representation of the humid region in ANN learning process. This approach can be further refined to standardize for minimum number of locations to be represented in generalized ANN models to produce adequate accuracy. This type of generalized ANN models can be used as an alternative to the existing conventional models.

### Comparison with regression modeling

In many respects, ANN models have the same basics of multiple regression modeling. For example the terms used in ANN modeling such as weight, input, output, example, training, bias, and hidden node correspond to coefficient, predictor, response, observation, parameter estimation, intercept, and derived predictor, respectively, in regression analysis. The ANNs are extremely flexible and have ability to represent wide range of responses and detect non-predefined relations such as nonlinear effects and/or interactions. The ANNs do not require standard regression assumptions of independence, normality and constancy of variance of residuals. Outliers in the response and predictors can have effect on models fit but the use of bounded logistic function tends to limit the influence of individual cases in comparison with regression. However, these advantages come at the cost of reduced interpretability of the model parameters/weights. The different elements of the weight matrix of ANN models resemble the parameters of regression modeling though in case of ANN models such parameters are significantly higher. Thus, theoretically, ANN should produce better result than regression model for same data length.

Evapotranspiration modeling using ANN has received much attention in the recent years. The methodology of ANN model development for evapotranspiration has proved its efficacy. The evapotranspiration modeling indeed is data intensive and requires us to handle huge datasets. This necessitates the use of high end computing techniques in evapotranspiration modeling. In the past few years, several researchers applied next generation computing techniques in evapotranspiration modeling those includes genetic algorithm (Kisi 2010; Kim and Kim 2008; Parsuraman et al. 2007), fuzzy-logic (Kisi and Ozturk 2007), and support vector machines (Kisi and Cimen 2009). These methodologies are, however, in the initial stages for development and can prove as the technique of future for evapotranspiration modeling.

### Conclusions

ANN modeling now has been added a new dimension in computational science. The application of ANN models in ETo estimation is now being widely discussed. Several contributions on ANN modeling in ETo estimation were reviewed in which all are validating the ANN as an alternate model to estimate ETo. Presently, the ANN modeling in ETo estimation has the characteristics of input–output (black-box) modeling which does not model the inherent physical process. Approach to develop the ANN models varied substantially by means of selection of best network architecture, learning scheme, and performance criteria. Most of the studies applied locally trained ANN models in ETo estimation. This has the limitation of global validity of such developed ANN models. However, some studies attempted to develop ANN models which have global validity, but these models need to be validated at several locations under varying sets of conditions in order to evaluate the generalization potential of ANNs in ETo modeling. This could be achieved either by developing ANN model to map ETo process both in space and time or by incorporating data from several locations to train the ANN.

### References

- ASCE Task Committee on Application of Neural Networks in Hydrology (2000a) Artificial neural network in hydrology. I: preliminary concepts. *J Hydrol Eng ASCE* 5(2):115–123
- ASCE Task Committee on Application of Neural Networks in Hydrology (2000b) Artificial neural network in hydrology. II: hydrologic application. *J Hydrol Eng ASCE* 5(2):124–137
- Berry MJA, Linoff G (1997) Data mining techniques. Wiley, NY
- Bruton JM, McClendon RW, Hoogenboom G (2000) Estimating daily pan evaporation with artificial neural networks. *Trans ASAE* 43(2):491–496
- Campolo M, Andreussi P, Sodalt A (1999) River stage forecasting with a neural network model. *Water Res Res* 35(4):1191–1197
- Cigizoglu HK (2003) Estimation, forecasting and extrapolation of flow data by artificial neural networks. *Hydro Sci J* 48(3):349–361
- Cochran WG, Cox GM (1957) Experimental designs, 2nd edn. Wiley, New York
- Dai X, Shi H, Li Y, Ouyang Z, Huo Z (2009) Artificial neural network models for estimating regional reference evapotranspiration based on climate factors. *Hydrol Process* 23(2009):442–450
- Dawson WC, Wilby R (1998) An artificial neural networks approach to rainfall-runoff modeling. *Hydrol Sci J* 43(1):47–66
- French MN, Krajewski WF, Cuykendall RR (1992) Rainfall forecasting in space and time using a neural network. *J Hydrol* 137:1–31
- Gomez KA, Gomez AA (1984) Statistical procedures in agricultural research, 2nd edn. Wiley, New York
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl Eng Agric* 1(2):96–99
- Hegazy T, Ayed A (1998) Neural network model for parametric cost estimation of highway projects. *J Constr Eng Manag ASCE* 124(3):210–218

- Jain SK, Nayak PC, Sudhir KP (2008) Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. *Hydrol Process* 22(13):2225–2234
- Jensen ME, Burman RD, Allen RG (1990) *Evapotranspiration and Irrigation Water Requirements*. ASCE Manual and Rep. on Engrg. Pract. No. 70. ASCE, New York
- Jones JW, Ritchie JT (1990) Crop growth models. In: Hoffman GJ, Howel TA, Solomon KH (eds) *Management of farm irrigation systems*. ASAE Monograph no. 9, ASAE, St. Joseph, MI, USA, pp 63–89
- Keskin ME, Terzi O (2006) Artificial neural network models of daily pan evaporation. *J Hydrol Eng* 11(1):65–70
- Khoob AR (2008a) Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semi-arid environment. *Irrig Sci* 26(3):253–259
- Khoob AR (2008b) Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. *Irrig Sci* 27(1):35–39
- Kim S, Kim HS (2008) Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *J Hydrol* 351:299–317
- Kisi O (2006) Generalized regression neural networks for evapotranspiration modeling. *Hydrol Sci J* 51(6):1092–1105
- Kisi O (2007) Evapotranspiration modeling from climatic data using a neural computing technique. *Hydrol Process* 21(14):1925–1934
- Kisi O (2008) The potential of different ANN techniques in evapotranspiration modeling. *Hydrol Process* 22:2449–2460
- Kisi O (2009) Daily pan evaporation modelling using multi-layer perceptrons and radial basis neural networks. *Hydrol Process* 23:213–223
- Kisi O (2010) Fuzzy genetic approach for modeling reference evapotranspiration. *J Irrig Drain Eng* 136(3):175–183
- Kisi O, Cimen M (2009) Evapotranspiration modeling using support vector machine. *Hydrol Sci J* 54(5):918–928
- Kisi O, Ozturk O (2007) Adaptive neuro-fuzzy computing technique for evapotranspiration modeling. *ASCE J Irrig Drain Eng* 133(4):368–379
- Kumar M, Raghuvanshi NS, Singh R, Wallender WW, Pruitt WO (2002) Estimating evapotranspiration using artificial neural network. *J Irrig Drain Eng ASCE* 128(4):224–233
- Kumar M, Bandyopadhyay A, Raghuvanshi NS, Singh R (2008) Comparative study of conventional and artificial neural network-based ETo estimation models. *Irrig Sci* 26(6):531–545
- Kumar M, Raghuvanshi NS, Singh R (2009) Development and validation of GANN model for evapotranspiration estimation. *J Hydrol Eng ASCE* 44(2):131–140
- Landeras G, Ortiz-Barredo A, Lopez JJ (2008) Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the basque country (Northern Spain). *Agric Water Manag* 95(5):553–565
- LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten ZIP code recognition. *Neural Comput* 1:541–551
- LeCun Y, Simard PY, Pearlmetter B (1993) Automatic Learning Rate Maximization by On-Line Estimation of the Hessian's Eigenvectors. In: Hanson SJ, Cowan JD, Giles CL (eds) *Advances in neural information processing systems* 5. Morgan Kaufmann, San Mateo, CA, pp 156–163
- Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and application. *Environ Model Softw* 15:101–124
- Markham IS, Rakes TR (1998) The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. *Comput Ops Res* 25(4):251–263
- Marti P, Royuela A, Manzano J, Palau-Salvador G (2010) Generalization of ETo ANN models through data supplanting. *J Irrig Drain Eng* 136(3):161–174
- Moghaddamnia A, Gousheh MG, Piri J, Amin S, Han D (2009) Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Adv Water Res* 32:88–97
- Monteith JL (1965) *Evaporation and Environment*. Proceedings of the state and movement of water in living organisms. XIXth Symposium, Soc. For Exp. Biol. Swansea, Cambridge University Press, New York, pp 205–234
- Odhiambo LO, Yoder RE, Yoder DC, Hines JW (2001) Optimization of fuzzy evapotranspiration model through neural training with input-output examples. *Trans ASAE* 44(6):1625–1633
- Orr GB, Leen TK (1997) Using Curvature Information for Fast Stochastic Search. In: Mozer MC, Jordan MI, Petsche T (eds) *Advances in neural information processing systems* 9. The MIT Press, Cambridge, pp 606–612
- Parsuraman K, Elshorbgby A, Carey SK (2007) Modelling the dynamics of the evapotranspiration process using genetic programming. *Hydro Sci* 52(3):563–578
- Penman HL (1948) Natural evaporation from open water, bare soil and grass. *Proc R Soc Lond* 193:120–146
- Rao V, Rao H (1996) *C++ neural networks and fuzzy logic*. BPB Publications, B-14, Connaught Place, New Delhi-110001, pp 380–381
- Rumelhart RE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. *Parallel distributed processing*, vol Vol 1. MIT Press, Cambridge, pp 318–362
- SNNS (1995) User manual, version 4.1, report no. 6/95. Institute for Parallel and Distributed High Performance Systems. University of Stuttgart, Germany
- Snyder R, Pruitt W (1985) Estimating reference evapotranspiration with hourly data, Chap. VII. In: R. Snyder et al. (eds) *California irrigation management information system final report*. University of California-Davis. Land, Air and Water Resources Paper no. 10013, California, USA
- Sudheer KP, Gosain AK, Rangan M, Saheb SM (2002) Modelling evaporation using an artificial neural network algorithm. *Hydrol Process* 16:3189–3202
- Sudheer KP, Gosain AK, Ramasastri KS (2003) Estimating actual evapotranspiration from limited data using neural computing technique. *J Irrig Drain Eng ASCE* 129(3):214–218
- Swingler K (1996) *Applying neural networks: a practical guide*. Academic Press, London
- Trajkovic S (2005) Temperature-based approaches for estimating reference evapotranspiration. *J Irrig Drain Eng ASCE* 131(4):316–323
- Trajkovic S (2009) Comparison of radial basis function networks and empirical equations for converting from pan evaporation to reference evapotranspiration. *Hydrol Process* 23(2009):874–880
- Trajkovic S, Stankovic M, Todorovic B (2000) Estimation of FAO Blaney–Criddle b factor by RBF network. *J Irrig Drain Eng ASCE* 126(4):268–270
- Trajkovic S, Todorovic B, Stankovic M (2003) Forecasting reference evapotranspiration by artificial neural networks. *J Irrig Drain Eng ASCE* 129(6):454–457
- Traore S, Wang Y, Kerh T (2010) Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian Zone. *Agric Water Manage* 97:707–714
- Yeh IC (1998) Quantity estimation of building with logarithm-neuron networks. *J Constr Eng Manage ASCE* 124(5):210–218
- Zanetti SS, Sousa EF, Oliveira VPS, Almeida FT, Bernardo S (2007) Estimating evapotranspiration using artificial neural network and minimum climatological data. *J Irrig Drain Eng* 133(2):83–89